AD-A068 940 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
POWER SERIES METHODS II. THE HEAT EQUATION.(U)
FEB 79 R D SMALL DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1924 NL

| OF |
AD
AO 68940

END
DATE
FILMED
6-79
DDC

MRC Technical Summary Report #1924

POWER SERIES METHODS II
- THE HEAT EQUATION

See 1473 in back.

Robert D. Small

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

February 1979

(Received January 31, 1979)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Research Council of Canada
Montreal Road
Ottawa, Ontario K1A OR6
Canada

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

POWER SERIES METHODS II - THE HEAT EQUATION

Robert D. Small

Technical Summary Report # 1924

February 1979

ABSTRACT

The power series method developed by the author [1] is applied to the heat equation. Highly accurate semi-discrete systems of equations in t and in x are generated and are made stable by proper choice of parameters. A totally discrete scheme is produced that represents arbitrarily high accuracy in both x and t. Stability analysis indicates that while arbitrary order in t may be stable, the order of accuracy in x is restricted to be less than 16 and certain geometrical restrictions on the step sizes must be met. Truncation errors are examined and a consistency condition is obtained that further restricts the step sizes. The scheme is shown to coincide with Keller's Box Scheme [4] in its lowest order.

AMS (MOS) Subject Classifications: 35A40, 65M05, 65M10.

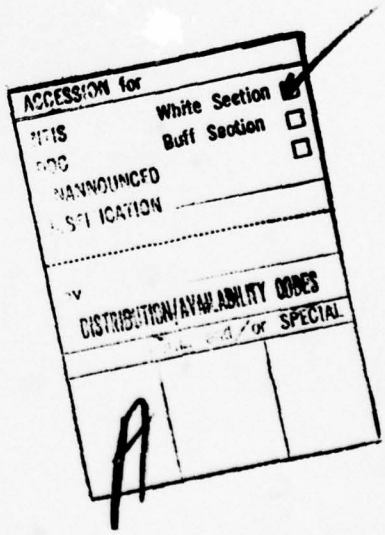
Key Words: Partial differential equation, heat equation, power series, difference scheme, high accuracy.

Work Unit Number 7 - Numerical Analysis.

Sponsored in part by the United States Army under Contract No. DAAG29-75-C-0024 and the National Research Council of Canada under Grant No. A8785.

SIGNIFICANCE AND EXPLANATION

In MRC TSR # 1923 the author introduced a method for solving partial differential equations by substitution of power series into the differential equation. In that report some technical properties of the method were discussed in connection with ordinary differential equations. This paper presents the first application of the method to a partial differential equation and takes as a basic example the heat equation. The method, in effect, substitutes a double power series with variables x and t, directly into the equation and finds the correct coefficients to satisfy the differential equation and the boundary conditions in a grid of cells. This simple procedure encounters many complications, however, and the bulk of the paper involves stepping slowly toward the goal of obtaining equations that govern the coefficients. Conditions on the built in parameters are derived that guarantee the errors due to the approximation to be small when introduced and not to grow with time.



POWER SERIES METHODS II - THE HEAT EQUATION

Robert D. Small

Introduction

In a previous paper [1], hereafter referred to as I, this author proposed a method whereby power series are used to generate arbitrarily accurate finite difference schemes for initial value problems in ordinary differential equations. It was shown there that a scheme may be made A-stable in the sense of Dalquist [2] by choosing the point of expansion of the power series to be the center of the interval in which the solution is sought. This paper, the sequel to I, applies the method to a particular partial differential equation, the heat equation, to illustrate in detail how a stable scheme with very high accuracy in both variables of a partial differential equation may be attained. Since we shall rely heavily upon the work of I we refer to equations of that paper with a preceding I and similarly a I precedes the figure number.

In a manner analogous to that of I, we fit power series solution surfaces over the domain that has been partitioned into rectangular cells. First infinite power series are found, and after initial and boundary conditions are satisfied the series are truncated. The truncations that lead to stable schemes are not as straightforward as in the case of ordinary differential equations. We find semi-discrete schemes, discretized in t in section 2 and in x in section 3. When these schemes are made stable by proper choice of some of the built in parameters, comparison of the two schemes leads to the totally discrete scheme of section 4. The stability analysis of this scheme then relates back to the analysis of I. Section 5 deals with truncation error and consistency and in section 6 we note briefly that the lowest order scheme coincides with Keller's Box Scheme.

Sponsored in part by the United States Army under Contract No. DAAG29-75-C-0024 and by the National Research Council of Canada under Grant No. A8785.

1. Set-up of the Problem

We now apply the power series method to the heat equation

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0 \quad . \tag{1}$$

The domain is partitioned into a grid of cells of length $2h$ in the x-direction and height $2k$ in the t-direction as illustrated in Figure 1. Each cell is identified by an ordered pair of integers (i,j) indicating its position in the grid. The solution or its approximation defined there is denoted $u_{ij}(x_i,t_j)$. Co-ordinates within each cell are denoted by (x_i,t_j) and are measured with respect to an origin that lies on the vertical mid-line of the cell a distance $2\lambda h$ from the bottom. They are related to the original ones by

$$x_i = x - (2i-1)h$$

$$t_j = t - 2(j-1+\lambda)k \quad .$$

This choice is made for comparison with the treatment of ordinary differential equations in I and λ will later be taken to be $1/2$ for reasons similar to those in I. For illustration we take a rectangular grid of L cells in width in the x-direction.

The solution u is prescribed as $U(x)$ initially, $g_1(t)$ on the left side of the domain and $g_2(t)$ on the right. Then if the domain is $0 \leq x \leq 2Lh$, $0 \leq t < \infty$, we can express all variables in local co-ordinates giving initial and boundary conditions as

$$\begin{aligned} u_{i1}(x_i,-2\lambda k) &= U_i(x_i) \qquad i = 1,2,\cdots,L \\ \left.\begin{aligned} U_{1j}(-h,t_j) &= g_{1j}(t_j) \\ U_{Lj}(h,t_j) &= g_{2j}(t_j) \end{aligned}\right\} & \; j = 1,2,\cdots,\infty \end{aligned} \tag{2}$$

where the quantities on the right side are defined by

$$U_i(x_i) = U(x),\ g_{1j}(t_j) = g_1(t),\ g_{2j}(t_j) = g_2(t) \quad .$$

Other domain shapes and boundary conditions, including free surfaces can be handled without major modification of this basic method.

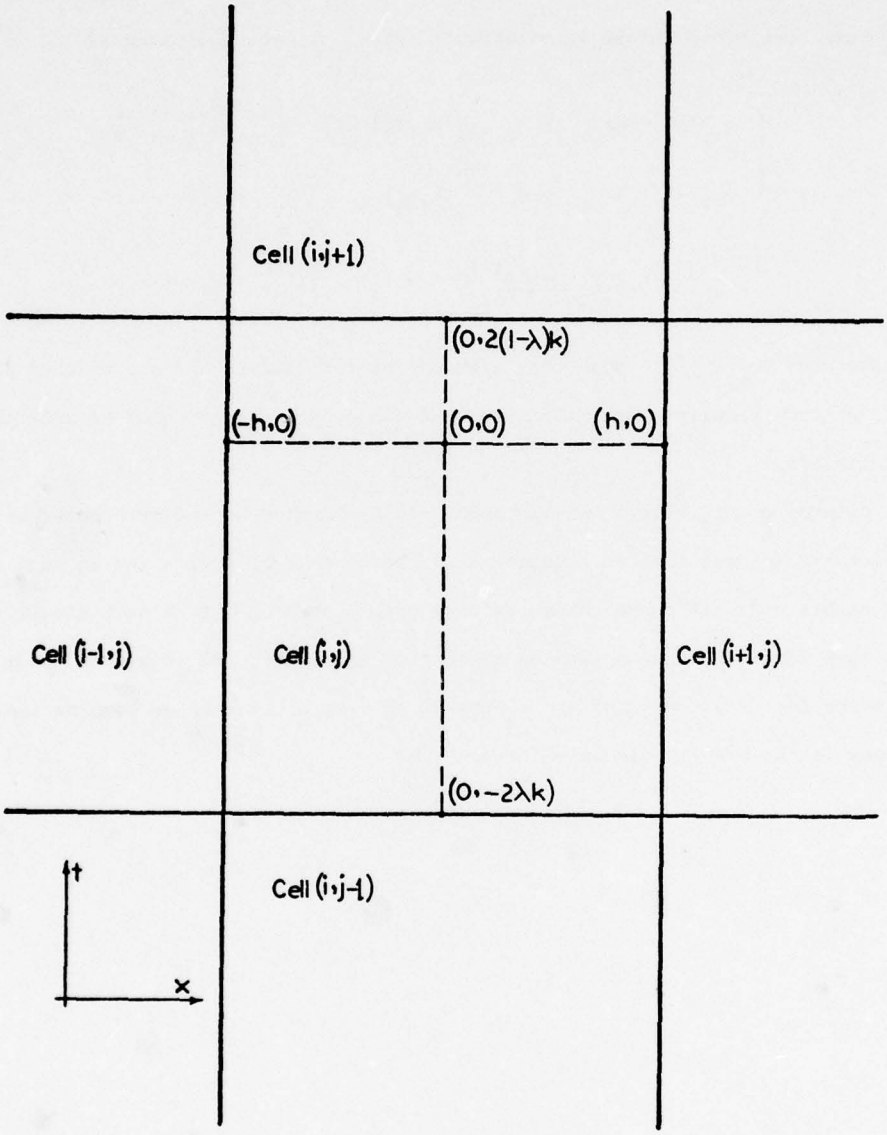


Figure 1. Typical Cell

For interior cell boundaries u is continuous across the horizontal boundaries and u and $\frac{\partial u}{\partial x}$ are continuous across vertical boundaries. These conditions are

$$
(3)\qquad \begin{cases} u_{ij}(x_i,-2\lambda k) = u_{i,j-1}(x_i,2(1-\lambda)k) & \begin{cases} i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \end{cases} \\ u_{ij}(-h,t_j) = u_{i-1,j}(h,t_j) \\ \dfrac{\partial u_{ij}}{\partial x_i}(-h,t_j) = \dfrac{\partial u_{i-1,j}}{\partial x_{i-1}}(h,t_j) \end{cases} \begin{cases} i = 2,3,\cdots,L \\ j = 1,2,\cdots,\infty \end{cases} .
$$

Commas are used to separate subscript expressions for clarity and are omitted when no ambiguity arises. No two subscripts are multiplied and subscripts are not used to indicate partial differentiation.

We determine $U_{ij}(x_i,t_j)$ within each cell by finding an infinite power series in x_i and t_j satisfying (1) and then we truncate it. The process of truncating in such a way as to leave a stable and well posed scheme is very complicated and it is much simplified by treating one variable first and then beginning again with the other. We encounter conditions that must be satisfied for each semi-discrete scheme to be stable and then we combine the two sets of conditions in the totally discrete formulation.

2. Discretization in t

Choosing to discretize in t first, we write u_{ij} as the power series in t_j,

$$u_{ij}(x_i,t_j) = \sum_{n=0}^{\infty} \frac{a_{ijn}(x_i)t_j^n}{n!} \quad .$$

Substituting this into (1) we obtain a difference equation for a_{ijn} which has the solution

$$a_{ijn}(x_i) = a_{ij}^{(2n)}(x_i)$$

for some function $a_{ij}(x_i)$ yet to be determined. A superscript (n) indicates the nth derivative of the function with respect to its argument. Then u_{ij} takes the form

$$u_{ij}(x_i,t_j) = \sum_{n=0}^{\infty} \frac{a_{ij}^{(2n)}(x_i)t_j^n}{n!} \quad . \tag{4}$$

It remains to select from (4) those solutions that satisfy boundary conditions (2) and (3). Substitution into (2) gives

$$\left.\begin{aligned} \sum_{n=0}^{\infty} \frac{a_{il}^{(2n)}(x_i)}{n!}(-2\lambda k)^n &= U_i(x_i) \qquad i = 1,2,\cdots,L \\ \sum_{n=0}^{\infty} \frac{a_{1j}^{(2n)}(-h)}{n!} t_j^n &= g_{1j}(t_j) \\ \sum_{n=0}^{\infty} \frac{a_{Lj}^{(2n)}(h)}{n!} t_j^n &= g_{2j}(t_j) \end{aligned}\right\} \quad j = 1,2,\cdots,\infty \quad . \tag{5}$$

We retain the first equation as it stands, but in the second and third we equate coefficients of like power of t_j to obtain

$$\left\{\begin{aligned} a_{1j}^{(2n)}(-h) &= g_{1j}^{(n)}(0) \\ a_{Lj}^{(2n)}(h) &= g_{2j}^{(n)}(0) \end{aligned}\right\} \left\{\begin{aligned} &j = 1,2,\cdots,\infty \\ &n = 0,1,\cdots,\infty \quad . \end{aligned}\right. \tag{6}$$

Similarly, substitution of (4) into (3) gives

$$(7)\qquad \sum_{n=0}^{\infty} \frac{a_{ij}^{(2n)}(x_i)}{n!}(-2\lambda k)^n = \sum_{n=0}^{\infty} \frac{a_{i,j-1}^{(2n)}(x_i)}{n!}(2(1-\lambda)k)^n \quad \begin{cases} i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \end{cases}$$

$$(8)\qquad a_{ij}^{(n)}(-h) = a_{i-1}^{(n)}(h) \quad \begin{cases} i = 2,3,\cdots,L \\ j = 2,3,\cdots,\infty \\ n = 0,1,\cdots,\infty \end{cases} .$$

We see now that, for fixed j and assuming that the $a_{i,j-1}$ are known functions, (7) is a coupled set of L infinite order ordinary differential equations in the L functions a_{ij} and there are the equivalent of L sets of boundary data given by (6) and (8). In addition, (5) is a set of ordinary differential equations that starts the whole process at $j = 1$. This is an unnecessarily complicated interpretation, however, since (8) states that a_{ij} is continuous to all orders of derivatives in the x-direction and hence we may disregard the subscript i and solve for one function a_j in a horizontal strip of cells at time step j that extends the entire width of the domain.

So far the problem is exact. We may approximate it by truncating the series in (5) and (7) at $n = N$ which leaves an order $2N$ ordinary differential equation for the one function a_{ij}. Then the proper boundary conditions selected from (6) are those that apply for $n = 0,1,\cdots,N-1$, and the continuity conditions (8) apply for the range $n = 0,1,\cdots,2N-1$. Having solved for all the a_{ij} one obtains the solution u_{ij} by substituting into (4) which has been truncated at $n = N$. Thus the semi-discrete scheme of order N, discretized in t, is

$$(9a)\qquad \sum_{n=0}^{N} \frac{a_{il}^{(2n)}(x_i)}{n!}(-2\lambda k)^n = U_i(x_i) \qquad i = 1,2,\cdots,L$$

$$(9b)\qquad \begin{cases} a_{lj}^{(2n)}(-h) = g_{lj}^{(n)}(0) \\ a_{Lj}^{(2n)}(h) = g_{2j}^{(n)}(0) \end{cases} \quad \begin{cases} j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,N-1 \end{cases}$$

$$(9c)\qquad \sum_{n=0}^{N} \frac{a_{ij}^{(2n)}(x_i)}{n!}(-2\lambda k)^n = \sum_{n=0}^{N} \frac{a_{i,j-1}^{(2n)}(x_i)}{n!}(2(1-\lambda)k)^n \quad \begin{cases} i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \end{cases}$$

$$(9d)\qquad a_{ij}^{(n)}(-h) = a_{i-1,j}^{(n)}(h) \quad \begin{cases} i = 2,3,\cdots,L \\ j = 2,3,\cdots,\infty \\ n = 0,1,\cdots,2N-1 \end{cases}$$

$$u_{ij}(x_i,t_j) = \sum_{n=0}^{N} \frac{a_{ij}^{(2n)}(x_i)}{n!} t_j^n \begin{cases} i = 1,2,\cdots,L \\ j = 1,2,\cdots,\infty \end{cases} . \tag{9e}$$

The problem is now discrete in the time direction, j having replaced t_j, but still continuous in space as x_i is present and i is superfluous. This system may be tested for stability using Von Neumann's criterion. Since (9) is a linear system, on an infinite domain the error satisfies the same equations as the solution and we can examine one Fourier component of the error by setting

$$a_{ij}(x_i) = \beta^j e^{\hat{i}\alpha x_i} \hat{a}$$

where α is any real number corresponding to the wave number of an error, $\hat{a}$ is any amplitude of error, β gives the rate of growth of the error, which is to be determined, and $\hat{i} = \sqrt{-1}$. Substituting this into (9c), we obtain, for $\hat{a} \neq 0$,

$$\sum_{n=0}^{N} \frac{[\beta\lambda^n - (\lambda-1)^n]}{n!} (2\alpha^2 k)^n = 0 . \tag{10}$$

Solving for β we have

$$\beta = \sum_{n=0}^{N} [2(\lambda-1)\alpha^2 k]^n / n! \Big/ \sum_{n=0}^{N} (2\lambda\alpha^2 k)^n / n! .$$

For α real, β is real and for the scheme to be stable we must have $|\beta| \leq 1$ so that the error does not grow. We shall have stability, therefore, if for $|\beta| > 1$, (10) has no real roots for α. This is the case if the term in square brackets has consistent sign for all n, which occurs when $\lambda \geq 1/2$. Thus the scheme (9) is stable if we choose $\lambda \geq 1/2$.

3. Discretization in x

We now discretize in the x-direction to obtain further stability conditions that must be met in the full discretization. The power series in x_i that satisfies (1) is

$$u_{ij}(x_i,t_j) = \sum_{m=0}^{\infty} \frac{\bar{a}_{ij}^{(m)}(t_j)}{(2m)!} x_i^{2m} + \sum_{m=0}^{\infty} \frac{\bar{b}_{ij}^{(m)}(t_j)}{(2m+1)!} x_i^{2m+1} \tag{11}$$

where the $\bar{a}_{ij}$ and the $\bar{b}_{ij}$ are functions of t_j yet to be determined. Substitution of (11) into (2) and (3) gives

$$\left.\begin{cases} \bar{a}_{il}^{(m)}(-2\lambda k) = U_i^{(2m)}(0) \\ \bar{b}_{il}^{(m)}(-2\lambda k) = U_i^{(2m+1)}(0) \end{cases}\right\} \begin{cases} i = 1,2,\cdots,L \\ m = 0,1,\cdots,\infty \end{cases} \tag{12a}$$

$$\left.\begin{cases} \sum_{m=0}^{\infty} \frac{\bar{a}_{1j}^{(m)}(t_j)}{(2m)!} h^{2m} - \sum_{m=0}^{\infty} \frac{\bar{b}_{1j}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} = g_{1j}(t_j) \\ \sum_{m=0}^{\infty} \frac{\bar{a}_{Lj}^{(m)}(t_j)}{(2m)!} h^{2m} + \sum_{m=0}^{\infty} \frac{\bar{b}_{Lj}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} = g_{2j}(t_j) \end{cases}\right\} j = 1,2,\cdots,\infty \tag{12b}$$

$$\left.\begin{cases} \bar{a}_{ij}^{(m)}(-2\lambda k) = \bar{a}_{i,j-1}^{(m)}(2(1-\lambda)k) \\ \bar{b}_{ij}^{(m)}(-2\lambda k) = \bar{b}_{i,j-1}^{(m)}(2(1-\lambda)k) \end{cases}\right\} \begin{matrix} m = 0,1,\cdots,\infty \\ i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \end{matrix} \tag{12c}$$

$$\left.\begin{aligned} &\sum_{m=0}^{\infty} \frac{\bar{a}_{ij}^{(m)}(t_j)}{(2m)!} h^{2m} - \sum_{m=0}^{\infty} \frac{\bar{b}_{ij}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} \\ &\qquad = \sum_{m=0}^{\infty} \frac{\bar{a}_{i-1,j}^{(m)}(t_j)}{(2m)!} h^{2m} + \sum_{m=0}^{\infty} \frac{\bar{b}_{i-1,j}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} \\ &-\sum_{m=0}^{\infty} \frac{\bar{a}_{ij}^{(m+1)}(t_j)}{(2m+1)!} h^{2m+1} + \sum_{m=0}^{\infty} \frac{\bar{b}_{ij}^{(m)}(t_j)}{(2m)!} h^{2m} \\ &\qquad = \sum_{m=0}^{\infty} \frac{\bar{a}_{i-1,j}^{(m+1)}(t_j)}{(2m+1)!} h^{2m+1} + \sum_{m=0}^{\infty} \frac{\bar{b}_{i-1,j}^{(m)}(t_j)}{(2m)!} h^{2m} \end{aligned}\right\} \begin{matrix} i = 2,3,\cdots,L \\ j = 1,2,\cdots,\infty \end{matrix} . \tag{12d}$$

We now interpret (12b) and (12d) as $2L$ coupled infinite order ordinary differential equations in the $2L$ functions $\bar{a}_{ij}$ and $\bar{b}_{ij}$ with $2L$ sets of data in (12a) for the case $j = 1$ and $2L$ sets of data in (12c) for $j \geq 2$. The continuity conditions (12c), however, indicate that both $\bar{a}_{ij}$ and $\bar{b}_{ij}$ are continuous to all orders of derivatives in the t-direction and hence j plays no role in the solutions. We are really solving for $2L$ functions in L vertical strips that extend infinitely forward in time.

We approximate this exact system by truncating the series in (12d) at finite values of m. Displaying only the summation signs, we modify (12d) to

$$(13) \qquad \begin{cases} \sum_{m=0}^{M_1} - \sum_{m=0}^{M_2} = \sum_{m=0}^{M_1} + \sum_{m=0}^{M_2} \\ - \sum_{m=0}^{M_3} + \sum_{m=0}^{M_4} = \sum_{m=0}^{M_3} + \sum_{m=0}^{M_4} \end{cases} .$$

Then both of equations (12b) are truncated as in the first of (13) so that equations that maintain continuity of the function value are of consistent accuracy.

Now the differential equations are discrete in space and continuous in time. We postpone selecting the correct boundary conditions from among (12a) and (12c) until the upper limits of (13) are chosen to make that system stable. We examine one Fourier component of the error by setting

$$(14) \qquad \begin{cases} \bar{a}_{ij}(t_j) = e^{\beta t_j} e^{2\hat{i}i\alpha h}\, \hat{a} \\ \bar{b}_{ij}(t_j) = e^{\beta t_j} e^{2\hat{i}i\alpha h}\, \hat{b} \end{cases}$$

for real arbitrary α, $\hat{a}$, $\hat{b}$. Substituting (14) into (13) we obtain two linear homogeneous equations in $\hat{a}$ and $\hat{b}$. The condition for non-trivial solutions is

$$(15) \qquad \sum_{m=0}^{M_1} \frac{(\beta h^2)^m}{(2m)!} \sum_{m=0}^{M_4} \frac{(\beta h^2)^m}{(2m)!} \left(1-e^{-2\hat{i}\alpha h}\right)^2 = \sum_{m=0}^{M_2} \frac{(\beta h^2)^{m+1/2}}{(2m+1)!} \sum_{m=0}^{M_3} \frac{(\beta h^2)^{m+1/2}}{(2m+1)!} \left(1+e^{-2\hat{i}\alpha h}\right)^2 .$$

The expression $\left(\frac{1+e^{-2\hat{i}\alpha h}}{1-e^{-2\hat{i}\alpha h}}\right)^2$ simplifies to $-\cot^2 \alpha h$.

Since α is arbitrary, this expression can take any negative value and we set it equal to $-c^2$. We can take square roots of (15) if we choose $M_1 = M_4$ and $M_2 = M_3$. Then we must have $M_1 = M_2$ or $M_1 = M_2 + 1$ for maximum accuracy since this does not retain in (13) any term more accurate than one omitted. Taking the square root of (15) and setting

$$\beta h^2 = z^2 \tag{16}$$

and

$$M_1 = [M/2] \quad , \quad M_2 = [(M-1)/2] \tag{17}$$

we obtain

$$\sum_{m=0}^{M_1} \frac{z^{2m}}{(2m)!} = \hat{i}c \sum_{m=0}^{M_2} \frac{z^{2m+1}}{(2m+1)!} \tag{18}$$

which is precisely equation (I-16) derived in I in connection with A-stability of this method and M plays the role of N of (I-16).

For stability we require that $\text{Re}\ \beta \leq 0$ in order that errors (14) do not grow with time. Accordingly, (16) indicates that we have stability if the roots of (18) satisfy

$$\frac{\pi}{4} \leq \arg z \leq \frac{3\pi}{4} \quad \text{or} \quad \frac{5\pi}{4} \leq \arg z \leq \frac{7\pi}{4} \quad . \tag{19}$$

The roots of equation (18) form closed curves parametrized by c. These curves originally appeared in Figure I-1 and are illustrated later in Figure 2b. Condition (19) is satisfied by these curves for $M \leq 15$ and hence $M = 15$ is the greatest order of accuracy that may be used stably.

We now choose the appropriate boundary conditions from (12a) and (12c) to suit differential equations (12b) and (12d). The semi-discrete scheme of order M, discretized in x, is

$$\left.\begin{array}{ll} \bar{a}_{il}^{(m)}(-2\lambda k) = U_i^{(2m)}(0) & m = 0,1,\cdots,M_2 \\ \bar{b}_{il}^{(m)}(-2\lambda k) = U_i^{(2m+1)}(0) & m = 0,1,\cdots,M_1-1 \end{array}\right\} \quad i = 1,2,\cdots,L \tag{20a}$$

(20b)
$$\left\{ \begin{array}{l} \sum_{m=0}^{M_1} \frac{\bar{a}_{1j}^{(m)}(t_j)}{(2m)!} h^{2m} - \sum_{m=0}^{M_2} \frac{\bar{b}_{1j}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} = g_{1j}(t_j) \\ \sum_{m=0}^{M_1} \frac{\bar{a}_{Lj}^{(m)}(t_j)}{(2m)!} h^{2m} + \sum_{m=0}^{M_2} \frac{\bar{b}_{Lj}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} = g_{2j}(t_j) \end{array} \right\} \quad j = 1,2,\cdots,\infty$$

(20c)
$$\left\{ \begin{array}{ll} \bar{a}_{ij}^{(m)}(-2\lambda k) = \bar{a}_{i,j-1}^{(m)}(2(1-\lambda)k) & m = 0,1,\cdots,M_2 \\ \bar{b}_{ij}^{(m)}(-2\lambda k) = \bar{b}_{i,j-1}^{(m)}(2(1-\lambda)k) & m = 0,1,\cdots,M_1-1 \end{array} \right\} \quad \begin{array}{l} i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \end{array}$$

(20d)
$$\left\{ \begin{array}{l} \sum_{m=0}^{M_1} \frac{\bar{a}_{ij}^{(m)}(t_j)}{(2m)!} h^{2m} - \sum_{m=0}^{M_2} \frac{\bar{b}_{ij}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} \\ \qquad = \sum_{m=0}^{M_1} \frac{\bar{a}_{i-1,j}^{(m)}(t_j)}{(2m)!} h^{2m} + \sum_{m=0}^{M_2} \frac{\bar{b}_{i-1,j}^{(m)}(t_j)}{(2m+1)!} h^{2m+1} \\ -\sum_{m=0}^{M_2} \frac{\bar{a}_{ij}^{(m+1)}(t_j)}{(2m+1)!} h^{2m+1} + \sum_{m=0}^{M_1} \frac{\bar{b}_{ij}^{(m)}(t_j)}{(2m)!} h^{2m} \\ \qquad = \sum_{m=0}^{M_2} \frac{\bar{a}_{i-1,j}^{(m+1)}(t_j)}{(2m+1)!} h^{2m+1} + \sum_{m=0}^{M_1} \frac{\bar{b}_{i-1,j}^{(m)}(t_j)}{(2m)!} h^{2m} \end{array} \right\} \quad \begin{array}{l} i = 2,3,\cdots,L \\ j = 1,2,\cdots,\infty \end{array}$$

(20e)
$$u_{ij}(x_i,t_j) = \sum_{m=0}^{M_1} \frac{\bar{a}_{ij}^{(m)}(t_j)}{(2m)!} x_i^{2m} + \sum_{m=0}^{M_2} \frac{\bar{b}_{ij}^{(m)}(t_j)}{(2m+1)!} x_i^{2m+1} \quad \left\{ \begin{array}{l} i = 1,2,\cdots,L \\ j = 1,2,\cdots,\infty \end{array} \right. .$$

4. Full Discretization

As indicated previously, the purpose of discretizing in both t and x separately is to determine the conditions that must be satisfied in the full discretization. We now take scheme (9) which is continuous in x_i and discretize it in the same manner that we made (1) discrete in x_i to produce (20).

We make the following substitution in (9),

$$a_{ij}(x_i) = \sum_{m=0}^{\infty} \frac{d_{ijm}}{m!} x_i^m \quad .$$

When powers of x_i arise, coefficients of equal powers of x_i are equated. The following set of equations is obtained.

(21a)
$$\sum_{n=0}^{N} \frac{d_{il,m+2n}}{n!} (-2\lambda k)^n = U_i^{(m)}(0) \quad \begin{cases} i = 1,2,\cdots,L \\ m = 0,1,\cdots,\infty \end{cases}$$

(21b)
$$\left\{ \begin{array}{l} \sum_{m=0}^{\infty} \frac{d_{1j,m+2n}}{m!} (-h)^m = g_{1j}^{(n)}(0) \\ \sum_{m=0}^{\infty} \frac{d_{Lj,m+2n}}{m!} h^m = g_{2j}^{(n)}(0) \end{array} \right\} \quad \begin{cases} j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,N-1 \end{cases}$$

(21c)
$$\sum_{n=0}^{N} \frac{d_{ij,m+2n}}{n!} (-2\lambda k)^n = \sum_{n=0}^{N} \frac{d_{i,j-1,m+2n}}{n!} (2(1-\lambda)k)^n \quad \begin{cases} i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \\ m = 0,1,\cdots,\infty \end{cases}$$

(21d)
$$\sum_{m=0}^{\infty} \frac{d_{ij,m+n}}{m!} (-h)^m = \sum_{m=0}^{\infty} \frac{d_{i-1,j,m+n}}{m!} h^m \quad \begin{cases} i = 2,3,\cdots,L \\ j = 2,3,\cdots,\infty \\ n = 0,1,\cdots,2N-1 \end{cases}$$

(21e)
$$u_{ij}(x_i,t_j) = \sum_{n=0}^{N} \sum_{m=0}^{\infty} \frac{d_{ij,m+2n}}{m!n!} x_i^m t_j^n \quad \begin{cases} i = 1,2,\cdots,L \\ j = 1,2,\cdots,\infty \end{cases}$$

In a similar fashion we discretize scheme (20) by the substitutions

$$\bar{a}_{ij}(t_j) = \sum_{n=0}^{\infty} \frac{\bar{c}_{ijn}}{n!} t_j^n \quad , \quad \bar{b}_{ij}(t_j) = \sum_{n=0}^{\infty} \frac{\bar{d}_{ijn}}{n!} t_j^n \quad ,$$

obtaining the set of equations

$$\text{(22a)}\qquad \left\{ \begin{array}{ll} \displaystyle\sum_{n=0}^{\infty} \frac{\bar{c}_{il,m+n}}{n!}(-2\lambda k)^n = U_i^{(2m)}(0) & m = 0,1,\cdots,M_2 \\ \displaystyle\sum_{n=0}^{\infty} \frac{\bar{d}_{il,m+n}}{n!}(-2\lambda k)^n = U_i^{(2m+1)}(0) & m = 0,1,\cdots,M_1-1 \end{array} \right\} \; i = 1,2,\cdots,L$$

$$\text{(22b)}\qquad \left\{ \begin{array}{l} \displaystyle\sum_{m=0}^{M_1} \frac{\bar{c}_{lj,m+n}}{(2m)!}h^{2m} - \sum_{m=0}^{M_2} \frac{\bar{d}_{lj,m+n}}{(2m+1)!}h^{2m+1} = g_{1j}^{(n)}(0) \\ \displaystyle\sum_{m=0}^{M_1} \frac{\bar{c}_{Lj,m+n}}{(2m)!}h^{2m} + \sum_{m=0}^{M_2} \frac{\bar{d}_{Lj,m+n}}{(2m+1)!}h^{2m+1} = g_{2j}^{(n)}(0) \end{array} \right\} \left\{ \begin{array}{l} j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,\infty \end{array} \right.$$

$$\text{(22c)}\qquad \left\{ \begin{array}{ll} \displaystyle\sum_{n=0}^{\infty} \frac{\bar{c}_{ij,m+n}}{n!}(-2\lambda k)^n = \sum_{n=0}^{\infty} \frac{\bar{c}_{i,j-1,m+n}}{n!}(2(1-\lambda)k)^n & m = 0,1,\cdots,M_2 \\ \displaystyle\sum_{n=0}^{\infty} \frac{\bar{d}_{ij,m+n}}{n!}(-2\lambda k)^n = \sum_{n=0}^{\infty} \frac{\bar{d}_{i,j-1,m+n}}{n!}(2(1-\lambda)k)^n & m = 0,1,\cdots,M_1-1 \end{array} \right\} \left\{ \begin{array}{l} i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \end{array} \right.$$

$$\text{(22d)}\qquad \left\{ \begin{array}{l} \displaystyle\sum_{m=0}^{M_1} \frac{\bar{c}_{ij,m+n}}{(2m)!}h^{2m} - \sum_{m=0}^{M_2} \frac{\bar{d}_{ij,m+n}}{(2m+1)!}h^{2m+1} \\ \qquad = \displaystyle\sum_{m=0}^{M_1} \frac{\bar{c}_{i-1,j,m+n}}{(2m)!}h^{2m} + \sum_{m=0}^{M_2} \frac{\bar{d}_{i-1,j,m+n}}{(2m+1)!}h^{2m+1} \\ -\displaystyle\sum_{m=0}^{M_2} \frac{\bar{c}_{ij,m+n+1}}{(2m+1)!}h^{2m+1} + \sum_{m=0}^{M_1} \frac{\bar{d}_{ij,m+n}}{(2m)!}h^{2m} \\ \qquad = \displaystyle\sum_{m=0}^{M_2} \frac{\bar{c}_{i-1,j,m+n+1}}{(2m+1)!}h^{2m+1} + \sum_{m=0}^{M_1} \frac{\bar{d}_{i-1,j,m+n}}{(2m)!}h^{2m} \end{array} \right\} \left\{ \begin{array}{l} i = 2,3,\cdots,L \\ j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,\infty \end{array} \right.$$

$$\text{(22e)}\qquad u_{ij}(x_i,t_j) = \sum_{m=0}^{M_1}\sum_{n=0}^{\infty} \frac{\bar{c}_{ij,m+n}}{(2m)!n!}x_i^{2m}t_j^n + \sum_{m=0}^{M_2}\sum_{n=0}^{\infty} \frac{\bar{d}_{ij,m+n}}{(2m+1)!n!}x_i^{2m+1}t_j^n \quad \left\{ \begin{array}{l} i = 1,2,\cdots,L \\ j = 1,2,\cdots,\infty \end{array} \right. .$$

Equations (21) express exactly the semi-discrete scheme continuous in x_i and equations (22) express exactly the semi-discrete scheme continuous in t_j. The former is stable for all N if $\lambda \geq 1/2$ and the latter is stable for $M \leq 15$, where M_1 and M_2 are related to M by (17). These become fully discrete schemes if the infinite sums are truncated but it is difficult to see how to do this stably on either scheme in isolation. Each can serve to show how to truncate the expansion in one of either x_i or t_j and it remains to select common variables for them so that both truncations can be applied simultaneously.

Comparing (21e) and (22e) we see that the expansion variables are related by

$$\bar{c}_{ij,n} = d_{ij,2n}, \quad \bar{d}_{ijn} = d_{ij,2n+1} \quad . \tag{23}$$

Placing these into (22) we obtain equations resembling (21) but with sums infinite in n and finite in m. From these two sets of equations we take all the finite limits and obtain the totally discrete scheme which follows.

$$\sum_{n=0}^{N} \frac{d_{il,m+2n}}{n!} (-2\lambda k)^n = U_i^{(m)}(0) \quad \begin{cases} m = 0,1,\cdots,M-1 \\ i = 1,2,\cdots,L \end{cases} \tag{24a}$$

$$\left.\begin{aligned} \sum_{m=0}^{M} \frac{d_{lj,m+2n}}{m!} (-h)^m &= g_{1j}^{(n)}(0) \\ \sum_{m=0}^{M} \frac{d_{Lj,m+2n}}{m!} h^m &= g_{2j}^{(n)}(0) \end{aligned}\right\} \begin{cases} j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,N-1 \end{cases} \tag{24b}$$

$$\sum_{n=0}^{N} \frac{d_{ij,m+2n}}{n!} (-2\lambda k)^n = \sum_{n=0}^{N} \frac{d_{i,j-1,m+2n}}{n!} (2(1-\lambda)k)^n \quad \begin{cases} i - 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \\ m = 0,1,\cdots,M-1 \end{cases} \tag{24c}$$

$$\sum_{m=0}^{M} \frac{d_{ij,m+n}}{m!} (-h)^m = \sum_{m=0}^{M} \frac{d_{i-1,j,m+n}}{m!} h^m \quad \begin{cases} i = 2,3,\cdots,L \\ j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,2N-1 \end{cases} \tag{24d}$$

$$u_{ij}(x_i,t_j) = \sum_{m=0}^{M} \sum_{n=0}^{N} \frac{d_{ij,m+2n}}{m!n!} x_i^m t_j^n - \frac{d_{ij,M+2N}}{M!N!} x_i^M t_j^N \quad \begin{cases} i = 1,2,\cdots,L \\ j = 1,2,\cdots,\infty \end{cases} \quad . \tag{24e}$$

The last formula is used only for displaying the solution after the d_{ijn} have been computed. The last term in (24e) appears because $d_{ij,M+2N}$ is not computed and must be deleted from the double sum.

We now have a finite difference scheme of order (M,N) for the heat equation (1). This scheme is now tested for stability by taking a Fourier component of the error,

$$d_{ijn} = \beta^j e^{2\hat{i}i\alpha h}\hat{d}_n \quad .$$

Substitution into (24c) and (24d) gives the homogeneous set of equations

$$\sum_{n=0}^{N} \frac{(-\lambda)^n - \beta^{-1}(1-\lambda)^n}{n!} (2k)^n \hat{d}_{m+2n} = 0 \qquad m = 0,1,\cdots,M-1$$

(25)

$$\sum_{m=0}^{M} \frac{(-1)^n - e^{-2\hat{i}\alpha h}}{m!} h^m \hat{d}_{m+n} = 0 \qquad n = 0,1,\cdots,2N-1 \quad .$$

The condition for non-trivial solutions is that the following matrix have zero determinant.

$$D = \begin{bmatrix}
\gamma_0 & 0 & \gamma_1 & 0 & \gamma_2 & 0 \cdots & \gamma_N & 0 & 0 \cdots & 0 \\
0 & \gamma_0 & 0 & \gamma_1 & 0 & \gamma_2 \cdots & 0 & \gamma_N & 0 \cdots & 0 \\
\vdots & & & & & & & & & \vdots \\
0 & 0 & 0 \cdots & \gamma_0 & 0 & \gamma_1 & 0 & \gamma_2 & 0 \cdots & \gamma_N \\
\delta_0 & \delta_1 & \delta_2 \cdots & \delta_M & 0 & 0 \cdots & & & & 0 \\
0 & \delta_0 & \delta_1 \cdots & \delta_{M-1} & \delta_M & 0 \cdots & & & & 0 \\
\vdots & & & & & & & & & \vdots \\
0 & 0 & 0 \cdots & \delta_0 & \delta_1 & \cdots & & & & \delta_M
\end{bmatrix} \begin{matrix} \uparrow \\ \downarrow \\ \text{M rows} \\ \uparrow \\ \downarrow \\ \text{2N rows} \end{matrix}$$

where $\gamma_n = \dfrac{(-\lambda)^n - \beta^{-1}(1-\lambda)^n}{n!} (2k)^n$

and $\delta_m = \dfrac{(-1)^m - e^{-2\hat{i}\alpha h}}{m!} h^m \quad .$

Matrix D is of the form of a Sylvester matrix and its determinant is called the resultant of the two polynomials

$$p_1(z) = \sum_{n=0}^{N} \gamma_n z^{2n} \quad \text{and} \quad p_2(z) = \sum_{m=0}^{M} \delta_m z^m \quad .$$

It may be shown that $\det D = 0$ if and only if $p_1(z)$ and $p_2(z)$ have at least one common zero. The classical proof (see Bôcher, [3], for example) begins with the assumption that $p_1(z)$ and $p_2(z)$ have a common factor and builds from the fact that there are no polynomials $f_1(z)$ and $f_2(z)$ such that

$$f_1(z)p_1(z) + f_2(z)p_2(z) = 1 \quad .$$

We present an alternative proof here since the classical one cannot be extended to treat the more complicated matrix that we encounter in section 5 in connection with truncation errors.

We note first that $\det D = 0$ is an algebraic equation of degree M in β for fixed α and hence there are M solutions for β. Further we note that the substitution

$$\hat{d}_n = z^n \qquad n = 0,1,\cdots,2N+M-1 \tag{26}$$

into (25) gives only two independent equations,

$$p_1(z) = 0 \quad , \quad p_2(z) = 0 \quad , \tag{27}$$

since all others differ from one of these only by a factor of a power of z. If equations (27) hold simultaneously then (25) are satisfied. It remains to prove that there are not further solutions alternative to (26). For fixed α, $p_2(z) = 0$ gives M solutions for z. Then since $p_1(z) = 0$ is linear in β, the substitution for z gives M solutions for β as required and no alternative solutions exist. Thus there are non-trivial solutions for the error if (27) holds.

System (24) is stable if (27) holds only for $|\beta| \leq 1$ and not otherwise. To investigate this condition we assume that $|\beta| > 1$ and try to show that the zeros of $p_1(z)$ and $p_2(z)$ are disjoint for all real α. Then since $\det D = 0$ is an algebraic equation for β, there must be solutions for β and they must satisfy $|\beta| \leq 1$.

We first examine the equation $p_1(z) = 0$. It happens to be a form of equation (I-12). To find the boundary between the stable and unstable regions we set $|\overline{R}| = 1$ in (I-12) and this

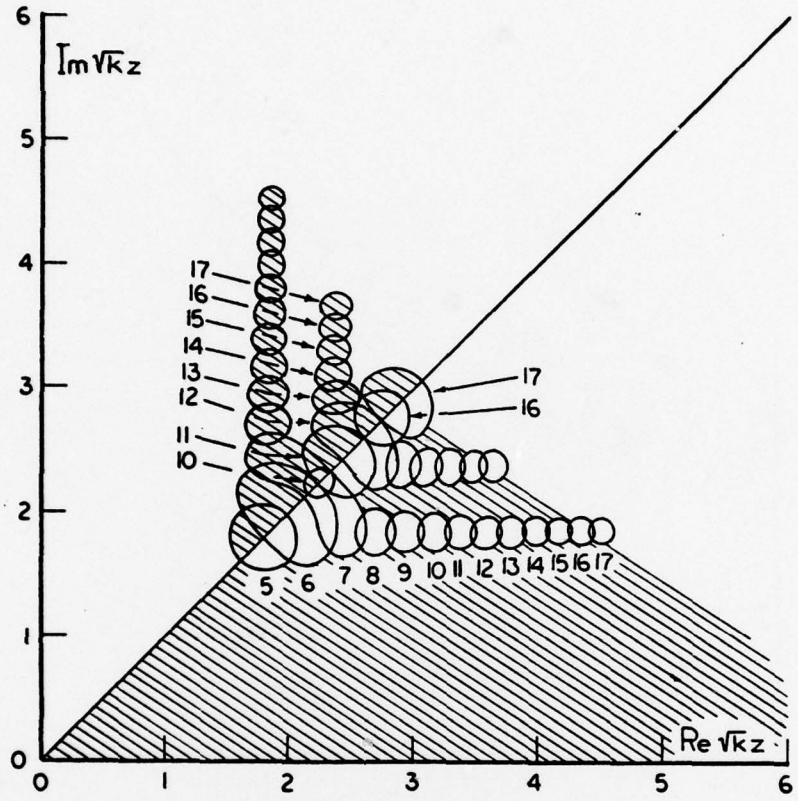


Figure 2a. Roots of $p_1(z) = 0$

The values of N are indicated. The regions where $|\beta| > 1$ are shaded.

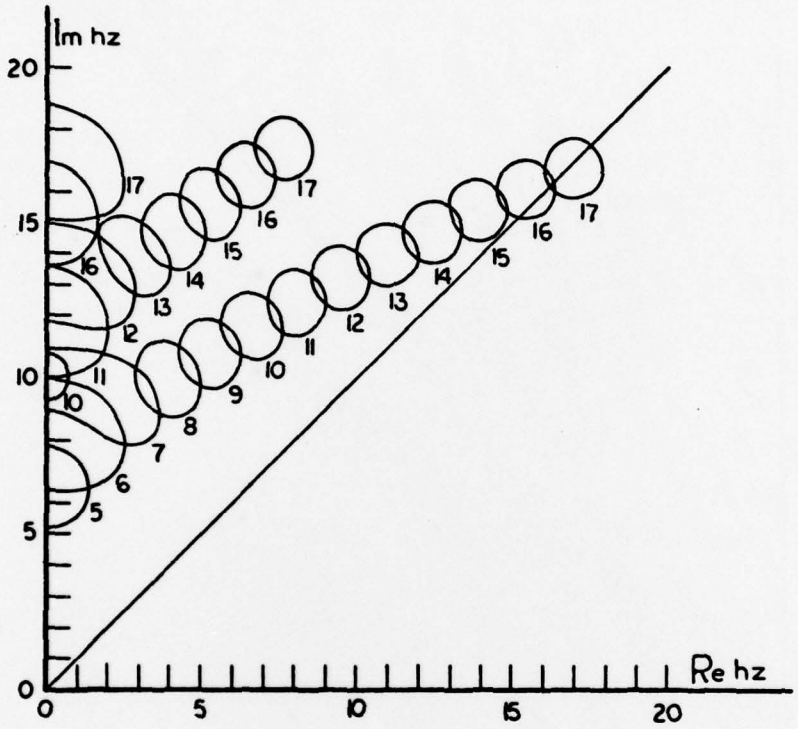


Figure 2b. Roots of $p_2(z) = 0$

The values of M are indicated.

we do by setting $\bar{R} = e^{i\phi}$ for real ϕ. The result is

$$\sum_{n=0}^{N} \frac{(1-\lambda)^n - e^{i\phi}(-\lambda)^n}{n!} (qh)^n = 0 \quad .$$

By comparison, if we wish to set $|\beta| = 1$ in $p_1(z) = 0$ we set $\beta = e^{i\phi}$ and obtain

$$\sum_{n=0}^{N} \frac{(1-\lambda)^n - e^{i\phi}(-\lambda)^n}{n!} (2kz^2)^n = 0 \quad . \tag{28}$$

The roots for qh found in I correspond to the roots for $2kz^2$. We found it necessary to set $\lambda = 1/2$ in order for inequality (I-14) to hold for $N \geq 3$ and this corresponds here to one of the conditions necessary to have $|\beta| \leq 1$ for all α when $N \geq 3$. Thus setting $\lambda = 1/2$, we have in Figure I-1 a plot of the roots of (28) where kz^2 is the variable. For comparison with the other polynomial, this plot is mapped into the plane with variable $\sqrt{k}\, z$ and its first quadrant is shown in Figure 2a. The regions where $|\beta| > 1$ appear shaded.

Now considering the equation $p_2(z) = 0$, if it is divided by $e^{-2\hat{i}\alpha h} - 1$ and the terms in α collected, we obtain

$$\sum_{m=0}^{M_1} \frac{(hz)^{2m}}{(2m)!} = \hat{i}c \sum_{m=0}^{M_2} \frac{(hz)^{2m+1}}{(2m+1)!} \tag{29}$$

where $c = \cot \alpha h$. We have seen this equation before as (I-16). Its roots for the variable hz, parameterized by c, form the closed curves and the imaginary axis of Figure I-1 and apply for orders M rather than N. The first quadrant of this plot is reproduced in Figure 2b.

The stability of a given scheme of order (M,N) depends on the positions of the curves corresponding to N in Figure 2a and M in Figure 2b. If any curve of Figure 2b intersects a shaded region of Figure 2a when the scales are made consistent, then the scheme is unstable. Clearly instability can be avoided if one of the rings of Figure 2a intersects with a ring of Figure 2b since the value of h or k can be adjusted to move the rings apart. Instability cannot be avoided when a ring of Figure 2b crosses a 45° line of Figure 2a and enters the large shaded region. This begins at $M = 16$. The rings of Figure 2a never touch the imaginary axis and thus arbitrarily large values of N may be used.

In summary, the fully discrete scheme (24) with $\lambda = 1/2$ is unconditionally stable for $1 \leq M \leq 4$, $1 \leq N \leq 4$. For $5 \leq M \leq 15$ and $5 \leq N < \infty$ there are mild restrictions on h and k for stability that are found graphically from Figures 2a and 2b. For $M \geq 16$ the scheme is unstable. These stability properties resemble those of the ordinary differential equation of I where, for $N \geq 5$ the step size h had to be adjusted so that the values of $\frac{qh}{2}$ avoided the small regions of Figure I-1.

5. Truncation Error and Consistency

A desirable attribute of a finite difference scheme is that arbitrarily high accuracy may be attained by allowing the step sizes to become small enough. While it neglects rounding error in the computer which becomes significant as many steps are taken, whether this feature holds is a gauge on the value of a difference scheme. The scheme is said to be consistent with the differential equation and its boundary conditions if the solution to the approximating difference equations converges to the exact solution of the differential equation as the step sizes vanish. The rate of convergence is determined by the truncation error which is the difference between approximate solution and the exact solution.

The exact solution $u_{ij}^{(e)}$ and its coefficients $d_{ijn}^{(e)}$ satisfy (24) with $M = \infty$, $N = \infty$ and $\lambda = 1/2$, whereas the computed solution u_{ij} and its coefficients d_{ijn} satisfy (24) for finite M and N and $\lambda = 1/2$. The $d_{ijn}^{(e)}$ are necessarily finite since the heat equation (1) is known to have analytic solutions. We let the truncation errors $\tilde{u}_{ij}$ and $\tilde{d}_{ijn}$ be defined by

$$\tilde{u}_{ij} = u_{ij} - u_{ij}^{(e)} \quad , \quad \tilde{d}_{ijn} = d_{ijn} - d_{ijn}^{(e)} \quad .$$

Subtraction of the two forms of (24) gives the following set of equations that governs the errors.

$$\text{(30a)} \qquad \sum_{n=0}^{N} \frac{\tilde{d}_{il,m+2n}}{n!} (-k)^n = \sum_{n=N+1}^{\infty} \frac{d_{il,m+2n}^{(e)}}{n!} (-k)^n \quad \begin{cases} i = 1,2,\cdots,L \\ m = 0,1,\cdots,M-1 \end{cases}$$

$$\text{(30b)} \qquad \left\{ \begin{array}{l} \displaystyle\sum_{m=0}^{M} \frac{\tilde{d}_{lj,m+2n}}{m!} (-h)^m = \sum_{m=M+1}^{\infty} \frac{d_{lj,m+2n}^{(e)}}{m!} (-h)^m \\ \\ \displaystyle\sum_{m=0}^{M} \frac{\tilde{d}_{Lj,m+2n}}{m!} h^m = \sum_{m=M+1}^{\infty} \frac{d_{Lj,m+2n}^{(e)}}{m!} h^m \end{array} \right\} \begin{cases} j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,N-1 \end{cases}$$

$$\text{(30c)} \qquad \sum_{n=0}^{N} \frac{\tilde{d}_{ij,m+2n}(-k)^n - \tilde{d}_{i,j-1,m+2n} k^n}{n!}$$

$$= \sum_{n=N+1}^{\infty} \frac{d_{ij,m+2n}^{(e)}(-k)^n - d_{i,j-1,m+2n}^{(e)} k^n}{n!} \quad \begin{cases} i = 1,2,\cdots,L \\ j = 2,3,\cdots,\infty \\ m = 0,1,\cdots,M-1 \end{cases}$$

(30d)
$$\sum_{m=0}^{M} \frac{\tilde{d}_{ij,m+n}(-h)^m - \tilde{d}_{i-1,j,m+n}\,h^m}{m!} = \sum_{m=M+1}^{\infty} \frac{d^{(e)}_{ij,m+n}(-h)^m - d^{(e)}_{i-1,j,m+n}\,h^m}{m!} \quad \begin{cases} i = 2,3,\cdots,L \\ j = 1,2,\cdots,\infty \\ n = 0,1,\cdots,2N-1 \end{cases}$$

(30e)
$$\tilde{u}_{ij}(x_i,t_j) = \sum_{m=0}^{M}\sum_{n=0}^{N} \frac{\tilde{d}_{ij,m+2n} + d^{(e)}_{ij,m+2n}}{m!n!} x_i^m t_j^n - \frac{(\tilde{d}_{ij,M+2N} + d^{(e)}_{ij,M+2N})}{M!N!} x_i^M t_j^N - \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{d^{(e)}_{ij,m+2n}}{m!n!} x_i^m t_j^n$$

$$i = 1,2,\cdots,L$$
$$j = 1,2,\cdots,\infty \quad .$$

Equations (30) yield solutions for the errors $\tilde{d}_{ijn}$ that are $O(h^{M+1},k^{N+1})$ if the coefficient matrix of the $\tilde{d}_{ijn}$ has a determinant that is not zero. If this is satisfied then the error in the solution $\tilde{u}_{ij}$ is $O(h^{M+1},k^{N+1},h^Mk^N)$ and the difference scheme (24) is consistent with (1), (2) and (3). Thus it remains to examine the homogeneous system.

In considering the homogeneous system we take the errors at the j-1 st time step to be small and investigate the error introduced due to advancing to the jth time step. The growth of errors already introduced is a matter of stability and was discussed in section 4. Deleting j, the homogeneous equations are, for all j,

(31a)
$$\sum_{n=0}^{N} \frac{\tilde{d}_{i,m+2n}}{n!}(-k)^n = 0 \quad \begin{cases} i = 1,2,\cdots,L \\ m = 0,1,\cdots,M-1 \end{cases}$$

(31b)
$$\left\{ \begin{array}{l} \sum_{m=0}^{M} \frac{\tilde{d}_{1,m+2n}}{m!}(-h)^m = 0 \\ \\ \sum_{m=0}^{M} \frac{\tilde{d}_{L,m+2n}}{m!}h^m = 0 \end{array} \right\} \quad n = 0,1,\cdots,N-1$$

(31c)
$$\sum_{m=0}^{M} \frac{\tilde{d}_{i,m+n}(-h)^m - \tilde{d}_{i-1,m+n}\,h^m}{m!} = 0 \quad \begin{cases} i = 2,3,\cdots,L \\ n = 0,1,\cdots,2N-1 \end{cases} \quad .$$

The coefficient matrix of (31) resembles the Sylvester matrix (25) in that a large number of rows are repeated with shifts to the right. This feature permits us to find solutions that contain powers of some number z that can be factored out of successive equations. It turns out that to obtain all of the independent solutions we must allow different coefficients for even and odd subscripts n of $\tilde{d}_{i,n}$. Thus we set

$$\left.\begin{aligned} \tilde{d}_{i,2n} &= A_i z^{2n} \qquad n = 0,1,\cdots,[N + \tfrac{M}{2} - \tfrac{1}{2}] \\ \tilde{d}_{i,2n+1} &= B_i z^{2n+1} \qquad n = 0,1,\cdots,[N + \tfrac{M}{2} - 1] \end{aligned}\right\} \quad i = 1,2,\cdots,L. \tag{32}$$

The independent equations after substitution into (31) are

$$\sum_{n=0}^{N} \frac{z^{2n}}{n!} (-k)^n = 0 \tag{33a}$$

$$\left\{\begin{array}{l} A_1 Z_1 - B_1 Z_2 = 0 \\ A_L Z_1 + B_L Z_2 = 0 \\ \\ A_i Z_1 - B_i Z_2 - A_{i-1} Z_1 - B_{i-1} Z_2 = 0 \\ -A_i Z_2 + B_i Z_1 - A_{i-1} Z_2 - B_{i-1} Z_1 = 0 \end{array}\right\} \quad i = 2,3,\cdots,L, \tag{33b}$$

where

$$Z_1 = \sum_{m=0}^{M_1} \frac{(zh)^{2m}}{(2m)!} \quad \text{and} \quad Z_2 = \sum_{m=0}^{M_2} \frac{(zh)^{2m+1}}{(2m+1)!}$$

and M_1 and M_2 are defined by (17).

There are non-trivial solutions to (31) if there are non-trivial solutions for A_i and B_i with z satisfying both (33a) and (33b). In turn, there are non-trivial solutions to (33b) if the $2L \times 2L$ matrix of coefficients is zero and this condition is

$$(34) \qquad \det = \begin{bmatrix} z_1 & -z_2 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ z_1 & z_2 & -z_1 & z_2 & 0 & 0 & \cdots & 0 & 0 \\ z_2 & z_1 & z_2 & -z_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & z_1 & z_2 & -z_1 & z_2 & \cdots & 0 & 0 \\ 0 & 0 & z_2 & z_1 & z_2 & -z_1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & & & & & \\ 0 & 0 & \cdots & 0 & 0 & z_2 & z_1 & z_2 & -z_1 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & z_1 & z_2 \end{bmatrix} = 0$$

Equations (33a) and (34), for a fixed choice of M, N and L, give a set of curves of k against h with parameter z that must be avoided when choosing h and k in order that the errors in (30) are of the orders stated. It would seem improbable that a random choice of h and k would satisfy these equations. In addition, these equations appear to be unrelated to the stability conditions of section 4, hence a choice of h and k that gives a stable scheme has a good chance of giving a consistent scheme. A choice of h and k that always produces a consistent scheme is found as follows. Choose h, and setting $z_1 = z_2$, solve for z and substitute into (33a) to get k. Then, while (33a) is satisfied, (34) is not since all the rows of the matrix are mutually orthogonal.

We have not yet shown that substitution (32) gives all the solutions to (31). This is a matter of checking that the number of solutions for k given h, determined by setting the determinant of the coefficient matrix of (31) equal to zero, is the same as the number of solutions for k resulting from solving (34) for z and substituting into (33a). This is a difficult task because the degree of equation (34) is not easily found in the general case. It has been checked in a large number of examples, however, and it appears that (32) does indeed lead to all solutions of (31).

6. Lowest Order Scheme and Concluding Remarks

If we set $M = 1$, $N = 1$, $\lambda = 1/2$ in (24) and eliminate the d_{ijn} in favor of the u_{ij} we obtain the following difference scheme, arranged to display the approximation of the derivatives of (1).

$$\frac{(u_{i+1,j} - u_{i+1,j-1}) + 2(u_{ij} - u_{i,j-1}) + (u_{i-1,j} - u_{i-1,j-1})}{4(2k)}$$

$$- \frac{(u_{i+1,j} - 2u_{i,j} + u_{i-1,j}) + (u_{i+1,j-1} - 2u_{i,j-1} + u_{i-1,j-1})}{2(2h)^2} = 0 \quad .$$

It has the Crank-Nicolson flavor, the approximation of the second derivative with respect to x being an average of second derivative formulas at the $j-1$ st and jth time steps and the approximation of the time derivative being a weighted average over the $i-1$ st, ith and $i+1$ st steps. This is in fact the Keller Box Scheme [4] for the heat equation and can be found by writing (1) as a first order system and using Keller's formulas for the first derivatives. In [4] Keller discusses, for parabolic equations that include the heat equation, A-stability, convergence, Richardson extrapolation and methods of solving the difference equations, some of which material will carry over to the higher order schemes. We speculate that the lowest order scheme in the power series method for $\lambda = 1/2$, applied to any partial differential equation, coincides with Keller's Box Scheme for that equation with a mild modification for problems involving free surfaces. We intend to show elsewhere that the power series method can treat irregular domains and free surface problems as a matter of course and that the lowest order scheme coincides with Keller's Box Scheme with an additional equation that adjusts the background grid of cells to account for the changing domain boundaries. In effect, a co-ordinate transformation is introduced automatically by the power series method that in other finite difference methods must be made before approximating the equations.

In conclusion, the power series method has been applied to the heat equation and a very accurate difference representation has been produced. The procedure followed here suggests that one should discretize in each variable separately and accumulate stability conditions. In this example discretization in t indicated that the point of expansion within each cell

should be on the vertical mid-line, above the center of the cell. Then any accuracy in t is acceptable for mild restrictions on the half time step k. Discretization in x indicated that the scheme is stable for orders of accuracy in x less than 16 for mild restrictions on the half space step h. The manner of truncation in each variable indicated how to obtain a well posed system when both the series are truncated simultaneously. The stability conditions under total discretization were modified to require that the point of expansion be precisely the center of the cell for mild restrictions on h and k, while the condition that $M \leq 15$ persisted. The "mild restrictions" just referred to are geometrical and derive from the root locus diagram of the key equation (I-16) that was involved in proving A-stability of the method. We shall show in a succeeding paper in this series by an analogous treatment of the wave equation that the stability conditions of the semi-discrete schemes may actually be more restrictive than for the totally discrete scheme. It turns out that one can obtain an arbitrarily accurate totally discrete scheme for the wave equation but not arbitrarily accurate semi-discrete schemes of the power series type. This treatment also depends heavily on equation (I-16). Finally, the condition that the scheme be consistent with the partial differential equation and the boundary conditions provides further restrictions on h and k, and when the scheme is consistent the truncation errors decrease as the orders of accuracy in x and t increase.

REFERENCES

[1] Small, R. D., "Power Series Methods I - Ordinary Differential Equations", MRC Technical Summary Report #1923, 1979.

[2] Dalquist, G., "A special stability problem for linear multistep methods", BIT 3, (1963), 27-43.

[3] Bôcher, M., "Introduction to higher algebra", Macmillan, 1938.

[4] Keller, H. B., "A New Difference Scheme for Parabolic Problems" in Numerical Solutions of Partial Differential Equations - II, Synspade 1970, B. Hubbard, Ed., Academic Press 1971.

RDS/jvs

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: #1924
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE *(and Subtitle)*: (6) POWER SERIES METHODS II - THE HEAT EQUATION

5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): (10) Robert D. Small

8. CONTRACT OR GRANT NUMBER(s): A8785; (15) DAAG29-75-C-0024

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 7 - Numerical Analysis

11. CONTROLLING OFFICE NAME AND ADDRESS: See Item 18.

12. REPORT DATE: (11) February 1979

13. NUMBER OF PAGES: 27

14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*: (14) MRC-TSR-1924

15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

(12) 32 p.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

(9) Technical summary rept.,

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Research Council of Canada
Montreal Road
Ottawa, Ontario K1A OR6
Canada

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Partial differential equation, heat equation, power series, difference scheme, high accuracy.

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The power series method ~~developed by the author~~ [1] is applied to the heat equation. Highly accurate semi-discrete systems of equations in t and in x are generated and are made stable by proper choice of parameters. A totally discrete scheme is produced that represents arbitrarily high accuracy in both x and t. Stability analysis indicates that while arbitrary order in t may be stable, the order of accuracy in x is restricted to be less than 16 and certain geometrical restrictions on the step sizes must be met. Truncation errors are examined and a consistency condition is obtained that further

→ next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

221 200

ABSTRACT (continued)

restricts the step sizes. The scheme is shown to coincide with Keller's Box scheme [4] in its lowest order.